

聴感印象の視覚化とサウンドデザイン

梶 友理香†
Yurika Kaji

カストナー マークアウエル††
Marc A. Kastner

石光 俊介††
Shunsuke Ishimitsu

†広島市立大学 情報科学部
††広島市立大学 情報科学研究科



1. 研究背景・研究目的



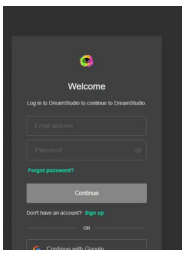
“スポーディで重厚感のあるエンジンサウンド”、“倍音成分が際立った澄んだエンジンサウンド”、“エレガントな中に重厚感があり、ボーカルが中央に定位すると共に楽音がきちんとナチュラルに分離して・・・”



イメージで呈示

2. Stable Diffusion

入力されたテキストに基づいて画像を生成
深層学習モデルを用いてテキストの意味を解析し、それに基づいて画像を生成
多段階の拡散プロセスを用いることで、テキストの意味を忠実に反映した画像を生成



3. 音響心理指標

ラウドネス (Loudness)

音の大きさ

定常音については ISO 5 3 2 Bで規格化

シャープネス (Sharpness)

甲高さ

高低域の音のバランスが高域側に偏ったときに感じる

ラフネス (Roughness)

粗さ感 ざらざら、ぶるぶる

ラウドネスが短い周期で変動する時に感じる (20~3000Hzまで)

変動強度

変動感 滑らかさの逆

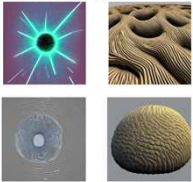
ラウドネスがゆっくりとした周期で変動するときに感じる (20Hzまで)

トナリティ

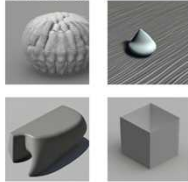
純音感, 変動のなさ

4. 音響心理指標の視覚化

ラウドネス強



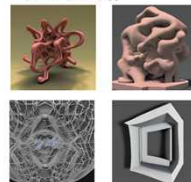
ラウドネス弱



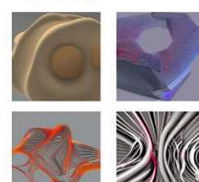
トナリティ強



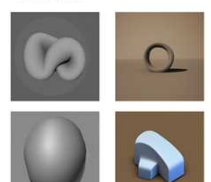
トナリティ弱



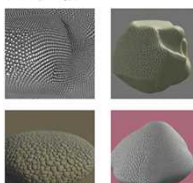
変動強度強



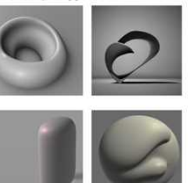
変動強度弱



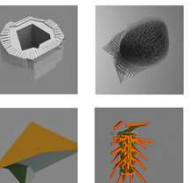
ラフネス強



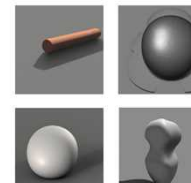
ラフネス弱



シャープネス強



シャープネス弱



トナリティや変動強度に関しては、期待される印象を十分に表現できていない。

トナリティに関しては音の色合いや質感を視覚的に再現が難
変動強度についても音の強さの変化を的確に反映する画像の生成に限界。

5. まとめ

文献

Stable Diffusion と Chat GPT を組み合わせることで、音の印象の視覚的表現を検討
音の印象を完全に可視化するには課題が残っており、特に正確性や精度の向上が課題
今後の研究では、実際のエンジン音を使い、適切な語句の選定と生成画像の分析を行い、
音の印象をより正確に可視化する技術の改善を目指す。

[1] Morris Alper & Hadar Averbuch-Elor Kiki or Boubu? Sound Symbolism in Vision-and-Language Models, NeurIPS 2023